



به نام آنکه جان را فکرت آموخت

وزارت بهداشت، درمان و آموزش پزشکی
معاونت آموزشی

دیرخانه شورای آموزش علوم پایه پزشکی، بهداشت و تخصصی
مرکز سنجش آموزش پزشکی

سوالات آزمون ورودی دکتری تخصصی (Ph.D)

سال تحصیلی ۱۴۰۵-۱۴۰۴

رشته: آمار زیستی

تعداد سوالات: ۱۳۰

زمان پاسخگویی: ۲۰۰ دقیقه

تعداد صفحات: ۲۱

مشخصات داوطلب:

نام:

نام خانوادگی:

توجه: استفاده از ماشین حساب مجاز نمی باشد.

پنج شنبه

۱۴۰۴/۰۸/۰۸

آمار زیستی





کلیات استنباط آماری

(استنباط آماری، تحلیل بقا و کارآزمایی بالینی)

۱- از یک توزیع نمایی $f(x) = \lambda e^{-\lambda x}$; $x, \lambda > 0$ فقط یک نمونه تصادفی انتخاب می‌شود. اگر ناحیه رد $R = \{X : X > k\}$ باشد، فرضیه $H_0 : \lambda = 1$ در مقابل فرضیه $H_1 : \lambda = 2$ رد می‌شود. برای خطای نوع اول ۰/۰۵ مقدار k چقدر است؟

(الف) $-Ln(0.05)$ (ب) $1 - Ln(0.05)$ (ج) $1 - Ln(0.95)$ (د) $-Ln(0.95)$

۲- در یک توزیع دوجمله‌ای با $n=2$ و پارامتر مجهول p ، جهت آزمون $H_0 : p = \frac{1}{3}$ در مقابل فرضیه $H_1 : p = \frac{2}{3}$ ، وقتی H_0 رد می‌شود که در یک تکرار آزمایش یک یا دو موفقیت مشاهده شده باشد. در این صورت، مقدار توان آزمون عبارتست از:

(الف) $\frac{5}{9}$ (ب) $\frac{4}{9}$ (ج) $\frac{8}{9}$ (د) $\frac{1}{9}$

۳- اگر $X_1, X_2, \dots, X_n$ نمونه‌ای تصادفی از توزیعی با پارامتر مجهول θ باشد، آماره $S(X)$ را یک آماره کمکی (Ancillary) می‌نامیم، هر گاه:

- (الف) توزیع آن مستقل از هر برآورد بسنده مینیمال θ باشد.
- (ب) توزیع آن به پارامتر مجهول θ وابسته نباشد.
- (ج) توزیع آن به توزیع نمونه‌ای برآورد پارامتر θ وابسته نباشد.
- (د) توزیع آن مستقل از هر برآورد بسنده θ باشد.

۴- اگر $X_1, X_2, \dots, X_n$ یک نمونه تصادفی، مستقل از هم و هم‌توزیع از توزیع کوشی با پارامترهای صفر و یک باشد، توزیع میانگین نمونه‌ای $(\bar{X})$ عبارتست از:

- (الف) نرمال استاندارد
- (ب) کوشی با پارامترهای صفر و n
- (ج) t با میانگین صفر و واریانس n
- (د) کوشی با پارامترهای صفر و یک

۵- اگر ۴۸ درصد مراجعین با سن بیش از ۶۵ سال به یک درمانگاه مرد و ۶۷ درصد مراجعین همین گروه سنی به پرفشاری خون مبتلا باشند؛ حداقل چند درصد این مراجعین، مرد و مبتلا به پرفشاری خون هستند؟

(الف) ۶۷ (ب) ۱۹ (ج) ۴۸ (د) ۱۵

۶- فرض کنید متغیر تصادفی X دارای چگالی احتمال $f(x) = \theta x^{\theta-1}$; $\theta > 0$, $0 < x < 1$ باشد که از آن فقط یک نمونه تصادفی با مقدار $\frac{1}{2}$ مشاهده شده باشد. برآورد حداکثر درست‌نمایی θ کدام است؟

(الف) $\frac{1}{2}$ (ب) $Ln2 - 1$ (ج) $Ln2$ (د) $\frac{1}{Ln2}$

۷- برای $0 \leq \alpha \leq 1$ آزمون با تابع توان $\beta(\theta)$ را در سطح α گوئیم، هرگاه:

الف) $Sup_{\theta \in \Theta_0} \beta(\theta) = \alpha$ ب) $Sup_{\theta \in \Theta_0} \beta(\theta) \leq \alpha$

ج) $Sup_{\theta \in \Theta_0} \beta(\theta) = \frac{\alpha}{2}$ د) $Sup_{\theta \in \Theta_0} \beta(\theta) \leq \frac{\alpha}{2}$

۸- اگر $X_1, X_2, \dots, X_n$ یک نمونه تصادفی، مستقل از هم و هم توزیع n تایی از توزیع برنولی با پارامتر p باشد، برآورد واریانس شانس موفقیت (Odds) در نمونه با کمک تقریب حاصل از بسط تیلور تقریباً برابر است با:

الف) $\frac{np}{(1-p)^3}$ ب) $\frac{np}{(1-p)^2}$ ج) $\frac{p}{n(1-p)^2}$ د) $\frac{p}{n(1-p)^3}$

۹- فرض کنید متغیر تصادفی X دارای توزیع دو جمله‌ای $X \sim Binomial(15, P)$ و P متغیر تصادفی دارای توزیع بتا با پارامترهای $\alpha = 1$ و $\beta = 2$ باشند؛ در این صورت $E(X)$ عبارتست از:

الف) ۲ ب) ۰/۵ ج) ۵ د) ۰/۲

۱۰- فرض کنید $X_1, X_2, \dots, X_{10}$ نمونه‌ای تصادفی، هم توزیع و مستقل از هم از توزیع یکنواخت پیوسته روی بازه صفر و θ باشد. در این صورت MSE برای برآوردگر $T = 2\bar{X}$ عبارتست از:

الف) $\frac{\theta}{15}$ ب) $\frac{\theta}{30}$ ج) $\frac{\theta^2}{30}$ د) $\frac{\theta^2}{15}$

۱۱- فرض کنید متغیر تصادفی X دارای توزیع پواسن با میانگین θ باشد. در این صورت، یک کران بالای چبیشف برای $P(X^2 \geq \theta)$ عبارتست از:

الف) $\theta^2 - 1$ ب) θ ج) $\theta + 1$ د) θ^2

۱۲- اگر واریانس متغیر تصادفی X مقداری برابر ۵ داشته باشد و مقدار مورد انتظار برای $Var(X|Y)$ عدد ۲ باشد؛ مقدار $Var[E(X|Y)]$ عبارتست از:

الف) ۴ ب) ۱ ج) ۲ د) ۳

۱۳- فرض کنید متغیر تصادفی X دارای توزیع نرمال با میانگین مجهول θ و واریانس ۱ باشد. در نمونه‌ای ۶ تایی تصادفی و مستقل از هم از این جامعه، میانگین نمونه ۱ بوده است. آماره نسبت درستنمایی در آزمون $H_0: \theta = 0$ در مقابل $H_1: \theta \neq 0$ کدام است؟

الف) $e^{-\frac{1}{6}}$ ب) $e^{-\frac{1}{3}}$ ج) e^{-3} د) e^{-6}

۱۴- فرض کنید $X_1, X_2, \dots, X_7$ نمونه‌ای تصادفی ۷ تایی، هم توزیع و مستقل از هم از توزیع یکنواخت پیوسته روی بازه‌ی صفر و یک باشد. در این صورت، انحراف معیار چهارمین آماره مرتب آن عبارتست از:

الف) $\frac{1}{25}$ ب) $\frac{1}{36}$ ج) $\frac{1}{5}$ د) $\frac{1}{6}$



۱۵- اگر $E[Y|x] = 2$ باشد، کدام گزینه در مورد $E[Y^2|x]$ صحیح است؟

(الف) $2 > E[Y^2|x] > 0$ (ب) $1 \geq E[Y^2|x] > 0$

(ج) $4 > E[Y^2|x] \geq 2$ (د) $E[Y^2|x] \geq 4$

۱۶- فرض کنید آماره آزمون دوطرفه میانگین تحت فرض صفر دارای توزیع نرمال استاندارد باشد. چنانچه مقدار مشاهده شده آماره آزمون برابر با ۲ باشد؛ در این صورت، مقدار احتمال (P-value) آزمون کدام است؟ (ϕ) بیان گر تابع توزیع نرمال استاندارد است)

(الف) $\phi(2)$ (ب) $1 - \phi(2)$ (ج) $2(1 - \phi(2))$ (د) $\phi(-2)$

۱۷- فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه‌ی تصادفی از یک توزیع برنولی با پارامتر نامعلوم p باشند. در این صورت، آماره‌ی بسنده و مینیمال برای p کدام است؟

(الف) $T(X) = \sum_{i=2}^n x_i$ (ب) $T(X) = \frac{\sum_{i=2}^n x_i}{n-1}$ (ج) $T(X) = \sum_{i=1}^n x_i$ (د) $T(X) = \frac{\sum_{i=2}^n x_i}{n}$

۱۸- فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه‌ی تصادفی دارای توزیع نمایی با چگالی $f(x) = \frac{1}{\theta} \exp\left(-\frac{x}{\theta}\right); \theta > 0, x > 0$ باشند. در این صورت، یک برآوردگر UMVUE برای θ^2 کدام است؟

(الف) $n\bar{X}^2$ (ب) $\frac{n+1}{n}\bar{X}^2$ (ج) $\frac{1}{n}\bar{X}^2$ (د) $\frac{n}{n+1}\bar{X}^2$

۱۹- فرض کنید X دارای توزیع پواسن با پارامتر θ و θ دارای توزیع گاما $\Gamma\left(\alpha, \frac{1}{\beta}\right)$ باشند. در این صورت، برآوردگر پسین بیزی θ کدام است؟

(الف) $\frac{x+\alpha}{\beta}$ (ب) $\frac{x+\alpha}{\beta+1}$ (ج) $\frac{\beta+1}{x+\alpha}$ (د) $(x+\alpha)(\beta+1)$

۲۰- فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه‌ی تصادفی دارای توزیع برنولی با پارامتر نامعلوم p باشند و $T(X) = \sum_{i=1}^n x_i$ تعریف شده باشد. در این صورت، یک برآوردگر UMVUE برای p^2 کدام است؟

(الف) $\bar{X}^2$ (ب) $\frac{T(T-1)}{n(n-1)}$ (ج) $\frac{T^2}{n}$ (د) $\frac{T(T-1)}{n^2}$

۲۱- فرض کنید تابع مولد گشتاور متغیر تصادفی X برابر با $M_X(t) = \exp(2t + 3t^2)$ باشد. در این صورت، واریانس X برابر است با:

(الف) ۴ (ب) ۶ (ج) ۳ (د) ۱۰

۲۲- فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه‌ی تصادفی n تایی از توزیعی با تابع چگالی

$$f(x) = \begin{cases} (\theta+1)x^\theta & 0 < x < 1; \theta > 0 \\ 0 & \text{otherwise} \end{cases}$$

باشد. در این صورت، برآورد گشتاوری پارامتر مجهول θ کدام است؟

- (الف) $1 - \bar{X}$ (ب) $1 - 2\bar{X}$ (ج) $\frac{1 - 2\bar{X}}{1 - \bar{X}}$ (د) $\frac{1 - 2\bar{X}}{\bar{X} - 1}$

۲۳- فرض کنید یک نمونه تصادفی سه تایی $\{-1, 0, 1\}$ از یک متغیر تصادفی X با جدول توزیع احتمال زیر مشاهده شده باشد. در این صورت، برآورد درستنمایی θ کدام است؟

X	-1	0	1
P(X=x)	$\frac{\theta}{4}$	$1 - \frac{\theta}{2}$	$\frac{\theta}{4}$

- (الف) $\frac{1}{4}$ (ب) $\frac{3}{4}$ (ج) $\frac{4}{3}$ (د) $\frac{3}{2}$

۲۴- فرض کنید θ شانس خط آمدن یک سکه باشد. در آزمون فرض $H_0: \theta = \frac{1}{2}$ در برابر $H_1: \theta = \frac{1}{3}$ اگر در پنج پرتاب مستقل، بیش از سه خط مشاهده شود؛ فرض صفر رد می‌شود. در این صورت، احتمال ارتکاب خطای نوع اول کدام است؟

- (الف) $\frac{5}{32}$ (ب) $\frac{1}{32}$ (ج) $\frac{6}{16}$ (د) $\frac{3}{16}$

۲۵- فرض کنید یک نمونه‌ی تصادفی سه تایی از توزیع برنولی با پارامتر θ در اختیار باشد. اگر تعداد پیروزی‌ها دست کم دو باشد، فرض $H_0: \theta = \frac{1}{2}$ رد می‌شود. در این صورت، احتمال ارتکاب خطای نوع دوم کدام است؟

- (الف) $3\theta^2 + 2\theta^3 + 1$ (ب) $3\theta^3 - 2\theta^2$ (ج) $2\theta^3 - 3\theta^2$ (د) $2\theta^3 - 3\theta^2 + 1$

۲۶- فرض کنید متغیر تصادفی X دارای توزیع نرمال با میانگین μ و واریانس ۹ باشد. نمونه‌ای ۹ تایی از این جامعه انتخاب و ناحیه‌ی بحرانی را به صورت $C = \{(X_1, X_2, \dots, X_9) | \bar{X} > k\}$ در نظر می‌گیریم. مقدار تقریبی k چند باشد

تا در آزمون فرض $\begin{cases} H_0: \mu = 1 \\ H_1: \mu = 2 \end{cases}$ مقدار α برابر ۰/۱ شود ($Z_{0.9} = 1.28$).

- (الف) $1/28$ (ب) $2/28$ (ج) $1/96$ (د) $2/32$

۲۷- فرض کنید X_1, X_2, X_3 یک نمونه تصادفی از توزیع یکنواخت در فاصله $(0,1)$ باشد. در این صورت، تابع چگالی میانه برابر است با:

- (الف) $6x(1-x)$ (ب) $3x(1-x)$ (ج) $6(1-x)^2$ (د) $3x^2$

۲۸- فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه تصادفی از توزیع نمایی با پارامتر λ باشد. چنانچه λ دارای توزیع پیشین گاما با پارامترهای (α, β) باشد؛ در این صورت، میانگین پسین λ برابر است با:

(د) $\frac{\alpha + n\bar{X}}{\beta + \bar{X}}$

(ج) $\frac{\alpha + \bar{X}}{\beta + n}$

(ب) $\frac{\alpha + n}{\beta + \bar{X}}$

(الف) $\frac{\alpha + n}{\beta + n\bar{X}}$

۲۹- فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه تصادفی از توزیعی با تابع چگالی $f(x, \theta) = \theta x^{\theta-1}$; $0 < x < 1, \theta > 0$ و همچنین $T = -\sum_{i=1}^n \log(x_i)$ باشد. در این صورت، $E(T)$ برابر است با:

(د) $\frac{n}{\theta}$

(ج) $\exp\left(-\frac{n}{\theta+1}\right)$

(ب) $\frac{n}{\theta+1}$

(الف) $\exp\left(-\frac{n}{\theta}\right)$

۳۰- فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه تصادفی از توزیع یکنواخت در فاصله $(0, \theta)$ باشد. در این صورت، یک فاصله اطمینان یک طرفه بالا با ضریب $1 - \alpha$ برای پارامتر θ برابر است با:

(د) $\theta \leq X_{(n)} \frac{1}{\alpha^n}$

(ج) $\theta \leq X_{(n)} - \frac{1}{\alpha^n}$

(ب) $\theta \leq X_{(n)} + \frac{1}{\alpha^n}$

(الف) $\theta \leq \frac{X_{(n)}}{(1-\alpha)^{\frac{1}{n}}}$

۳۱- اگر تابع بقا برابر با $S(t) = \frac{1}{1+2t}$; $t \geq 0$ باشد؛ در این صورت، تابع خطر برابر است با:

(د) $\frac{2t}{(1+2t)^2}$

(ج) $\frac{t}{1+2t}$

(ب) $\frac{2}{(1+t)^2}$

(الف) $\frac{2}{1+2t}$

۳۲- چنانچه زمان‌های بقا برابر با مقادیر $2^+, 2, 3^+, 5, 6^+, 6$ باشند؛ در این صورت، برآورد کاپلان مایر تابع بقا در زمان ۵ برابر است با:

(د) $\frac{3}{9}$

(ج) $\frac{6}{9}$

(ب) $\frac{5}{9}$

(الف) $\frac{4}{9}$

۳۳- در محاسبه آماره لگاریتم-رتبه‌ای برای مقایسه بقا در دو گروه، فرض کنید در زمان t ، تعداد افراد در معرض خطر در گروه ۱ برابر با ۱۰ نفر و تعداد افراد در معرض خطر در گروه ۲ برابر با ۲۰ نفر باشد. چنانچه در کل ۳ پیشامد در زمان t داشته باشیم که ۲ نفر آن در گروه ۱ باشد، تعداد رخداد مورد انتظار در زمان t در گروه ۱ برابر است با:

(د) $1/5$

(ج) $1/2$

(ب) ۱

(الف) ۲

۳۴- در مدل مخاطره متناسب کاکس، اگر تبدیل یکنوا روی زمان‌های بقا به کار رود؛ کدام یک از مقادیر زیر تغییر می‌کند؟

(الف) مقادیر ضرایب رگرسیون

(ب) مقدار تابع درستنمایی جزئی

(ج) مقدار معنی‌داری مدل

(د) هیچکدام

۳۵- برای یک مجموعه از داده‌های بقا، مدل شماره ۱ به صورت $h_0(t) \exp(\beta_1 x_1)$ و مدل شماره ۲ به صورت $h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2)$ است. چنانچه مقادیر لگاریتم درست‌نمایی مدل شماره ۱ و ۲ به ترتیب برابر با ۲۱۵- و ۲۱۲- باشد؛ در این صورت، مقدار آماره کای اسکور برای مقایسه دو مدل و نتیجه آزمون در سطح خطای ۵ درصد کدام است؟ (چندک ۹۵ درصد کای اسکور با درجه آزادی یک را برابر با ۴ در نظر بگیرید).

- الف) ۶، رد فرض صفر
- ب) ۳، عدم رد فرض صفر
- ج) ۶، عدم رد فرض صفر
- د) ۳، رد فرض صفر

۳۶- برای یک مدل کاکس با یک متغیر کمکی دوحالتی، پیشامدها در دو زمان t_1 و t_2 رخ داده است، اگر سهم درست‌نمایی جزئی به ترتیب برابر با $L_1 = 0.4$ و $L_2 = 0.3$ است. در این صورت، مقدار درست‌نمایی جزئی کل چقدر است؟

- الف) ۰.۷
- ب) $\text{Log}(0.3) + \text{Log}(0.4)$
- ج) $\text{Log}(0.4) \times \text{Log}(0.3)$
- د) ۰.۱۲

۳۷- در یک مدل کاکس که بیماران پرخطر را در مقابله با بیماران کم‌خطر مقایسه می‌کند (پرخطر با کد ۱، کم‌خطر با کد ۰)، برآورد نسبت خطر برابر با ۱/۵ (یک و نیم) به دست آمده است. در این صورت، تفسیر درست کدام است؟

- الف) بیماران پرخطر ۵۰٪ نرخ پیشامد آنی بالاتری دارند.
- ب) بیماران پرخطر ۱/۵ برابر بیشتر از بیماران کم‌خطر زنده می‌مانند.
- ج) در گروه پرخطر ۱/۵ پیشامد بیشتر به ازای هر بیمار رخ می‌دهد.
- د) میانگین زمان بقا در بیماران پرخطر به نصف کاهش می‌یابد.

۳۸- پس از برازش مدل کاکس با متغیرهای کمکی سن و درمان، باقیمانده‌های شوئنفلد متغیر درمان را استخراج کرده و آن‌ها را در برابر زمان رسم می‌کنیم. چنانچه ضریب همبستگی باقیمانده‌ها با زمان برابر با ۰/۲۸ و مقدار معنی‌داری برابر با ۰/۰۲ را نشان دهد؛ چه نتیجه‌ای می‌گیرید؟

- الف) خطرات متناسب برای درمان برقرار است ($P(PH) > 0.01$)
- ب) خطرات متناسب برای درمان برقرار است ($P(PH) > 0.1$)
- ج) خطرات متناسب برای درمان برقرار نیست ($P(PH) < 0.05$)
- د) خطرات متناسب برای درمان برقرار نیست ($P(PH) < 0.01$)

۳۹- در یک مدل رگرسیون شکست شتابیده (AFT) وایبول، برآورد نسبت زمان بقا برای متغیر دوسطحی درمان (درمان با کد ۱ و کنترل با کد ۰) برابر با ۲ شده است؛ در این صورت:

- الف) خطر وقوع رخداد برای گروه درمان $\exp(2)$ برابر خطر در گروه کنترل است.
- ب) زمان بقای مورد انتظار برای گروه درمان ۲ برابر زمان بقای گروه کنترل است.
- ج) لگاریتم زمان بقای گروه درمان ۲ برابر لگاریتم زمان بقای کنترل است.
- د) میانه زمان بقای گروه درمان برابر با ۵۰٪ میانه زمان بقای گروه کنترل است.

- ۴۰- در یک مدل رگرسیون شکست شتابیده (AFT) لگ لجستیک، برآورد ضریب رگرسیون متغیر دوسطحی درمان (درمان با کد ۱ و کنترل با کد ۰) برابر با 0.5 (نیم) شده است؛ در این صورت:
- الف) خطر وقوع رخداد برای گروه درمان $\exp(0.5)$ برابر خطر در گروه کنترل است.
- ب) لگاریتم زمان بقای گروه درمان ۲ برابر لگاریتم زمان بقای کنترل است.
- ج) زمان بقای مورد انتظار برای گروه درمان $\exp(0.5)$ برابر زمان بقای گروه کنترل است.
- د) میانه زمان بقای گروه درمان ۲ برابر میانه زمان بقای گروه کنترل است.

- ۴۱- در رویکرد حاشیه‌ای پیشامدهای بازگشتی، مجموعه در معرض خطر برای سومین عود چگونه تعریف می‌شود؟
- الف) فقط افرادی که دقیقاً دو پیشامد قبلی را تجربه کرده‌اند.
- ب) فقط افرادی که کمتر از سه پیشامد را تجربه کرده‌اند.
- ج) صرف‌نظر از تعداد پیشامدهای قبلی، همه افرادی که هنوز تحت پیگیری برای رخداد پیشامد سوم هستند.
- د) فقط افرادی که بیشتر از سه پیشامد را تجربه کرده‌اند.

- ۴۲- فرض کنید در مدل تعمیم یافته کاکس برآورد ضریب متغیر $x \times \log(\text{time})$ برابر با -0.6 با خطای استاندارد 0.2 باشد. چنانچه x یک متغیر تشدیدکننده بیماری بوده و ضریب آن در مدل مثبت و معنی‌دار باشد؛ در این صورت با فرض سطح معنی‌داری 5% ، کدام یک از گزینه‌های زیر درست است:
- الف) با افزایش زمان بقا، اثر متغیر x بر مخاطره تضعیف می‌گردد.
- ب) با افزایش زمان بقا، اثر متغیر x بر تابع بقا کاهشی است.
- ج) با افزایش زمان بقا، اثر متغیر x بر مخاطره تشدید می‌گردد.
- د) با افزایش زمان بقا، اثر متغیر x بر مخاطره ثابت است.

- ۴۳- در مبحث ریسک‌های رقابتی، در مدل فاین و گری، ضریب رگرسیون کدامیک از اثرات زیر را برآورد می‌کند؟
- الف) خطر اختصاصی علتی
- ب) خطر زیرتوزیع
- ج) خطر کل
- د) میانگین تابع تجمعی خطر

- ۴۴- فرض کنید $S_1(t)$ بیان‌گر تابع بقا برای افراد مبتلا به سرطان معده و $S_2(t)$ بیان‌گر تابع بقا افراد مبتلا به سرطان کولون باشد، چنانچه $S_1(t) = S_2(1.5 \times t)$ باشد، در این صورت؟
- الف) احتمال بقای یک فرد مبتلا به سرطان کولون ۶۰ ساله، برابر احتمال بقای یک فرد مبتلا به سرطان معده ۴۰ ساله است.
- ب) احتمال بقای یک فرد مبتلا به سرطان کولون ۴۰ ساله، برابر احتمال بقای یک فرد مبتلا به سرطان معده ۶۰ ساله است.
- ج) مخاطره‌ی یک فرد مبتلا به سرطان معده در زمان t ، $1/5$ برابر مخاطره‌ی یک فرد مبتلا به سرطان کولون در زمان t است.
- د) طول عمر افراد مبتلا به سرطان معده، 75% طول عمر افراد مبتلا به سرطان کولون است.

- ۴۵- فرض کنید تابع چگالی و تابع مخاطره‌ی بیماران مبتلا به سرطان مری به ترتیب به صورت $f(t) = \lambda p t^{p-1} e^{-\lambda t^p}$ و $h(t) = \lambda p t^{p-1}$ باشد؛ در این صورت، صدک ۲۵ام زمان بقا کدام است؟

الف) $\sqrt[p]{\frac{\ln(0.75)}{\lambda}}$ ب) $\sqrt[p]{\frac{\ln(0.25)}{\lambda}}$ ج) $\sqrt[p]{\frac{-\ln(0.25)}{\lambda}}$ د) $\sqrt[p]{\frac{-\ln(0.75)}{\lambda}}$

۴۶- یک کارآزمایی بالینی که دارای نقطه پایانی چندگانه (Multiple end-points) شامل چهار نقطه پایانی (End-point) می باشد. اگر برای هر نقطه پایانی یک آزمون آماری مجزا با سطح معنی داری (α) یکسان اجرا شود، آنگاه ممکن است کدام پیشامد روی دهد؟

- الف) خطر تولید نتیجه مثبت حقیقی زیاد می شود.
- ب) خطر تولید نتیجه منفی کاذب زیاد می شود.
- ج) خطر تولید نتیجه مثبت کاذب زیاد می شود.
- د) خطر تولید نتیجه منفی حقیقی زیاد می شود.

۴۷- در طراحی یک کارآزمایی بالینی، کدامیک از شرایط زیر بیشترین نقش را در تضمین اعتبار درونی (Internal validity) مطالعه ایفا می کند؟

- الف) انتخاب نمونه تصادفی از جامعه بیماران
- ب) انتخاب شرکت کنندگان از مراکز درمانی متعدد
- ج) تخصیص تصادفی شرکت کنندگان به گروه های مداخله و کنترل
- د) استفاده از معیارهای خروج سخت گیرانه برای حذف بیماران با بیماری های همراه

۴۸- در یک کارآزمایی بالینی دو سوکور (Double-blind) برای ارزیابی اثربخشی داروی ضدالتهاب، پژوهشگران متوجه شدند که برخی بیماران به دلیل عوارض جانبی خاص، احتمالاً متوجه نوع درمان خود شده اند. در این وضعیت کدام مورد زیر ایجاد می شود؟

- الف) افزایش توان آماری مطالعه به دلیل کاهش حجم نمونه
- ب) افزایش سوگیری مشاهده گر و تحلیل گر به دلیل افشای غیر مستقیم درمان
- ج) کاهش تعمیم پذیری نتایج به جمعیت عمومی
- د) کاهش احتمال خطای نوع دوم به دلیل عدم رعایت تصادفی سازی

۴۹- در طراحی یک کارآزمایی بالینی برای مقایسه دو درمان، پژوهشگران تصمیم می گیرند سطح معنی داری (α) را ۰/۰۵ و توان آزمون (Power) را ۹۰ درصد در نظر بگیرند. اگر تفاوت واقعی بین دو درمان کمتر از مقدار پیش بینی شده باشد، کدامیک از پیامدهای زیر محتمل تر است؟

- الف) کاهش توان آزمون و افزایش احتمال عدم کشف تفاوت واقعی
- ب) افزایش احتمال خطای نوع اول و رد نادرست فرض صفر
- ج) افزایش تعمیم پذیری نتایج به جمعیت های مختلف
- د) کاهش اعتبار درونی مطالعه

۵۰- در یک کارآزمایی بالینی، تفاوت میانگین فشار خون بین دو گروه درمانی برابر با ۲ میلی متر جیوه گزارش شده است و آزمون آماری نشان دهنده معناداری در سطح ($P < 0.01$) است. کدامیک از تحلیل های زیر در تفسیر نتایج مناسب تر است؟

- الف) نتیجه گیری قطعی درباره اثربخشی درمان به دلیل معناداری آماری
- ب) رد کامل درمان رقیب به دلیل تفاوت آماری مشاهده شده
- ج) بررسی اهمیت بالینی تفاوت ۲ میلی متر جیوه و معنی داری آماری آن در تصمیم گیری
- د) رد درمان جدید بدلیل ناچیز بودن اثر مشاهده شده

- ۵۱- در تعریف کارآزمایی بالینی منفی (Negative clinical trial) محقق به دنبال آنست که نشان دهد که اثر بخشی یک داروی جدید محافظه کارانه در مقایسه با اثر بخشی یک داروی استاندارد تهاجمی تر:
- الف) پایین تر است. ب) هم تراز نیست. ج) برتر است. د) هم تراز است.
- ۵۲- در انتشار نتایج کارآزمایی بالینی، کدامیک از موارد زیر می تواند منجر به سوگیری تفسیر شود؟
- الف) استفاده از تحلیل های چندگانه برای کنترل عوامل مخدوشگر
ب) تاکید و تصمیم بر اساس نتایج مقایسه مداخلات در زیر گروه ها
ج) توجه به اهمیت بالینی و معنی داری آماری نتایج
د) کاهش دادن سطح معنی داری
- ۵۳- اگر محقق بخواهد در طرح متقاطع (Crossover)، اثر انتقالی (Carryover) را با دقت بیشتر بررسی نماید آنگاه در طراحی مطالعه خود کدام حالت را باید مد نظر داشته باشد؟
- الف) تعداد تیمارها (Treatment) بیشتر شود.
ب) تعداد دنباله ها (Sequence) بیشتر شود.
ج) تعداد دوره ها (Period) بیشتر شود.
د) زمان Washout دو برابر گردد.
- ۵۴- در یک کارآزمایی بالینی، استفاده از تصادفی سازی بلوکی (Block randomization) با اندازه بلوک ثابت چه پیامدی برای مطالعه دارد؟
- الف) افزایش تعمیم پذیری ب) احتمال سوگیری ج) افزایش توان آماری د) کاهش توان آماری
- ۵۵- کدامیک از گزینه های زیر تعریف دقیق یک مطالعه دو سوکور (Double blind) است؟
- الف) بیمار و همراه بیمار از نوع درمان آگاهی ندارند.
ب) بیمار در دو مرحله دو نوع درمان را دریافت می کند.
ج) بیمار و درمانگر از نوع درمان که در حال اجرا است آگاهی ندارند.
د) بیمار از نوع درمان در حال اجرا آگاه بوده و درمانگر نوع درمان در حال اجرا را نمی داند.
- ۵۶- هدف اصلی کارآزمایی بالینی مرحله دوم عبارت است از:
- الف) تعیین سودمندی دارو یا روش درمانی بر روی بیماران داوطلب
ب) تعیین ایمنی دارو یا روش درمانی بر روی داوطلبان سالم
ج) تعیین سودمندی دارو یا روش درمانی بر روی داوطلبان سالم
د) تعیین تعداد نمونه در بررسی اثر دارو یا روش درمانی بر روی داوطلبان سالم
- ۵۷- رعایت اخلاق در انجام کارآزمایی بالینی مرحله سوم یعنی:
- الف) فقط التزام به رعایت اخلاق فردی (بیمار) بر اساس نظر درمانگر
ب) ایجاد تعادل در رعایت اخلاق فردی و جمعی (جامعه)
ج) فقط التزام به رعایت اخلاق جمعی (جامعه) بر اساس نظر درمانگر
د) نمونه گیری تصادفی از بیماران و تصادفی سازی آنها به گروه ها
- ۵۸- اگر بخواهیم تصادفی سازی را با کنترل جنس و تعداد غده های سینه افراد در کارآزمایی های بالینی دو گروهی در درمان سرطان سینه انجام دهیم کدامیک از گزینه زیر مناسب تر است؟
- الف) تصادفی سازی ساده (Simple randomization)
ب) تصادفی سازی بلوکی (Block randomization)
ج) تصادفی سازی طبقه ای (Stratified randomization)
د) تصادفی سازی بلوکی یا تصادفی سازی ساده

۵۹- کارآزمایی بالینی متقاطع (Crossover) دو گروهی دو دوره‌ای در مورد کدامیک از بیماری‌های زیر قابل اجرا است؟

- الف) آسم ، دیابت ، شکستگی پا
- ب) دیابت ، شکستگی پا، فشارخون
- ج) آسم ، دیابت ، فشارخون
- د) دیابت ، فشارخون ، سرما خوردگی

۶۰- اگر زمان مرگ ، فشار خون و شدت درد را در یک مطالعه کارآزمایی بالینی جمع آوری کنیم از نظر صحت کدام ترتیب برای داده‌ها مناسب‌تر است؟

- الف) اندازه‌گیری‌ها- داده‌های واقعی- داده‌های ارزشیابی بالینی
- ب) داده‌های ارزشیابی بالینی- اندازه‌گیری‌ها- داده‌های واقعی
- ج) داده‌های واقعی- داده‌های ارزشیابی بالینی- عقیده بیمار
- د) داده‌های واقعی- اندازه‌گیری‌ها- عقیده بیمار

روش‌های آمار زیستی

۶۱- در رگرسیون خطی، هدف از مرکزی کردن (centering) متغیرهای مستقل چیست؟

- الف) استانداردسازی
- ب) حذف مشاهدات دورافتاده
- ج) کاهش هم‌خطی بین متغیرهای مستقل
- د) ثبات واریانس

۶۲- کدام یک از گزینه‌های زیر درباره انحراف باقیمانده‌ها (residual deviance) در مدل‌های خطی تعمیم یافته درست است؟

- الف) هرچه نسبت residual deviance به درجه آزادی نزدیک به یک باشد، مدل برازش بهتری دارد.
- ب) هرچه نسبت residual deviance به درجه آزادی نزدیک به یک باشد، مدل دارای بیش برازش است.
- ج) هرچه residual deviance بزرگ‌تر باشد مدل برازش بهتری دارد.
- د) هرچه اختلاف residual deviance با درجه آزادی نزدیک به صفر باشد، مدل دارای بیش برازش است.

۶۳- در یک مدل خطی تعمیم یافته لجستیک با پارامتر p ، در چه حالتی تفاوت نتایج آزمون Wald و آزمون نسبت درست‌نمایی چشم‌گیر است؟

- الف) وقتی اندازه نمونه خیلی بزرگ باشد.
- ب) وقتی پارامتر مدل نزدیک یک یا صفر باشد.
- ج) همیشه نتایج یکسان هستند.
- د) وقتی پارامتر مدل نزدیک نیم باشد.

۶۴- کدامیک از موارد زیر از ویژگی‌های مهم توزیع نرمال چندمتغیره $\mathcal{N}_p(\mu, \Sigma)$ نیست؟

- الف) هر زیرمجموعه‌ای از متغیرها نیز دارای توزیع نرمال چندمتغیره است.
- ب) هر ترکیب خطی از متغیرها دارای توزیع نرمال تک‌متغیره است.
- ج) کوواریانس بین هر دو متغیر حتماً صفر است.
- د) تابع چگالی آن تابعی از فاصله مایلانوبیس از میانگین‌ها است.

۶۵- کدام مورد از مفروضات اساسی آزمون Hotelling's T^2 است؟

(الف) داده‌ها دارای توزیع نرمال چندمتغیره باشند.

(ب) متغیرها کاملاً مستقل باشند.

(ج) ماتریس واریانس-کوواریانس منفرد (singular) باشد.

(د) توزیع هر یک از متغیرها نرمال باشد.

۶۶- در تحلیل خوشه‌ای سلسله‌مراتبی با متغیر پاسخ کمی، کدام معیار فاصله رایج‌تر است؟

(الف) فاصله منهتن (ب) فاصله اقلیدسی (ج) فاصله آماری (د) فاصله کسینوسی

۶۷- در تحلیل مؤلفه‌های اصلی، مؤلفه‌های استخراج شده:

(الف) ترکیب غیرخطی از داده‌ها هستند.

(ب) همبسته‌اند.

(ج) کل واریانس داده‌ها را بیان می‌کنند.

(د) متعامد هستند.

۶۸- ماتریس واریانس - کوواریانس زیر مربوط به یک توزیع نرمال دو متغیره است:

$$\Sigma = \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix}$$

واریانس تعمیم یافته این توزیع چقدر است؟

(الف) ۱۰

(ب) ۱۲

(ج) ۴

(د) ۸

۶۹- در یک مطالعه، ۶۰ بیمار بر اساس واکنش به درمان به سه گروه تقسیم شده‌اند: خوب، متوسط، ضعیف. فراوانی‌های

مشاهده شده به ترتیب ۲۵، ۲۰، ۱۵ بوده است. برای بررسی یکنواخت بودن توزیع فراوانی‌ها آماره آزمون برابر است

با:

(الف) ۵

(ب) ۲,۵

(ج) ۰,۵

(د) ۵۰

۷۰- در مدل گرسیون خطی $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ برآوردگر حداقل مربعات معمولی β_0 و β_1 برابر است با:

$$\hat{\beta}_1 = \frac{\sum_i (Y_i - \bar{Y})^2}{\sum_i (X_i - \bar{X})^2}, \quad \hat{\beta}_0 = \left(\sum_i Y_i - \hat{\beta}_1 \sum_i X_i \right)^2 \quad (\text{الف})$$

$$\hat{\beta}_1 = \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_i (X_i - \bar{X})^2}, \quad \hat{\beta}_0 = \frac{1}{n} \left(\sum_i Y_i - \hat{\beta}_1 \sum_i X_i \right) \quad (\text{ب})$$

$$\hat{\beta}_1 = \frac{\sum_i (Y_i - \bar{Y})^2}{\sum_i (X_i - \bar{X})^2}, \quad \hat{\beta}_0 = \frac{1}{n} \left(\sum_i Y_i - \hat{\beta}_1 \sum_i X_i \right) \quad (\text{ج})$$

$$\hat{\beta}_1 = \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_i (X_i - \bar{X})^2}, \quad \hat{\beta}_0 = \left(\sum_i Y_i - \hat{\beta}_1 \sum_i X_i \right)^2 \quad (\text{د})$$

۷۱- در گرسیون چند گانه $Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon_i$ ، آماره آزمون t دو طرفه $H_0: \beta_2 = 0$ کدام گزینه زیر است؟

ب) $t = \frac{\hat{\beta}_2}{\sqrt{\hat{\sigma}^2 (X'X)^{-1}_{22}}}$ با درجه آزادی $n-2$

الف) $t = \frac{\hat{\beta}_2}{\sqrt{\hat{\sigma}^2 (X'X)^{-1}_{22}}}$ با درجه آزادی $n-3$

د) $t = \frac{\hat{\beta}_2}{\sqrt{\hat{\sigma}^2 (X'X)^{-1}_{22}}}$ با درجه آزادی $n-2$

ج) $t = \frac{\hat{\beta}_2}{\sqrt{\hat{\sigma}^2 (X'X)^{-1}_{22}}}$ با درجه آزادی $n-3$

۷۲- فاصله اطمینان $100(1-\alpha)\%$ برای میانگین پاسخ در x_0 در رگرسیون خطی ساده برابر کدام گزینه است؟

ب) $\hat{y}_0 \pm t_{1-\frac{\alpha}{2}, n-2} \hat{\sigma} \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{X})^2}{\sum_i (X_i - \bar{X})^2}}$

الف) $\hat{y}_0 \pm t_{1-\frac{\alpha}{2}, n-2} \hat{\sigma} \sqrt{1 + \frac{(x_0 - \bar{X})^2}{\sum_i (X_i - \bar{X})^2}}$

د) $\hat{y}_0 \pm z_{1-\frac{\alpha}{2}} \hat{\sigma} \sqrt{1 + \frac{(x_0 - \bar{X})^2}{\sum_i (X_i - \bar{X})^2}}$

ج) $\hat{y}_0 \pm z_{1-\frac{\alpha}{2}} \hat{\sigma} \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{X})^2}{\sum_i (X_i - \bar{X})^2}}$

۷۳- کدام گزینه در مورد نمودار باقیمانده‌ها در مدل رگرسیون خطی درست است؟

- الف) یک نمودار باقیمانده قیفی شکل (funnel) نشان دهنده غیرخطی بودن تابع میانگین است.
- ب) منحنی شکل بودن نمودار باقیمانده‌ها در مقابل مقادیر برازش شده، نشان دهنده ناهمگنی واریانس است.
- ج) وجود الگو در نمودار باقیمانده‌ها در مقابل مقادیر برازش شده، نشان دهنده اشتباه بودن مدل است.
- د) پراکندگی بدون الگو در باقیمانده‌ها، نشان دهنده حذف اشتباه یک متغیر مستقل است.

۷۴- کدام گزینه در مورد ماتریس $H = X(X'X)^{-1}X'$ صحیح است؟

- الف) متقارن و خودتوان است.
- ب) همواره یک ماتریس قطری با مولفه‌های قطری یکسان است.
- ج) اثر (Trace) ماتریس همیشه برابر با حجم نمونه است.
- د) همواره معکوس پذیر است.

۷۵- در یک مدل تحلیل واریانس با n تکرار و واریانس خطای σ^2 ، و دو عامل تصادفی A (با a سطح) و B (با b سطح) و اثر متقابل AB، امید ریاضی MS برای عامل A عبارت است از:

ب) $\sigma^2 + n\sigma_{AB}^2 + nb\sigma_A^2$

الف) $\sigma^2 + a\sigma_{AB}^2 + nb\sigma_A^2$

د) $\sigma^2 + a\sigma_{AB}^2 + ab\sigma_A^2$

ج) $\sigma^2 + n\sigma_{AB}^2 + ab\sigma_A^2$

۷۶- در یک مدل رگرسیون خطی چندگانه، کدام یک از گزینه‌های زیر صحیح است؟

- (الف) Leverage فقط بر اساس مقدار y اندازه‌گیری می‌شود.
 (ب) Cook's Distance میزان تأثیر حذف یک مشاهده بر روی تمام ضرایب را اندازه‌گیری می‌کند.
 (ج) آزمون Breusch-Pagan به منظور آزمودن خود همبستگی در مدل‌های رگرسیون خطی استفاده می‌شود.
 (د) در مقایسه دو مدل غیر لانه‌ای (non-nested)، آزمون F کلاسیک معیار مناسبی است.

۷۷- در رگرسیون خطی چندگانه $E(Y|X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$ اگر یکی از X ها را جایگزین Y کنیم:

- (الف) ضریب تعیین مدل تغییری نمی‌کند.
 (ب) می‌توان شاخص تورم واریانس را برای آن X محاسبه کرد.
 (ج) همخطی از بین می‌رود.
 (د) ضرایب رگرسیونی معکوس می‌شوند.

۷۸- برای مدل $y = 5x + 3$ با تعداد نمونه ۲۵ مقدار $var(\hat{\beta}) = 4$ به دست آمده است. فاصله اطمینان ۹۵ درصد برای

ضریب رگرسیونی تقریباً عبارت است از:

- (الف) (۷ و ۳) (ب) (۹ و ۱) (ج) (۵٫۸ و ۴٫۲) (د) (۵٫۴ و ۴٫۶)

۷۹- در رگرسیون لجستیک با متغیر پاسخ بیماری قلبی عروقی، سن به صورت کمی وارد مدل شده است. اگر ضریب

رگرسیونی سن $\beta = 0.05$ به دست آمده باشد نسبت شانس یک دهه افزایش در سن برای ابتلا به بیماری قلبی عروقی عبارت خواهد بود از:

- (الف) $Exp(0.05)$ (ب) $Exp(10)$ (ج) $Exp(0.5)$ (د) $Exp(-0.05)$

۸۰- مجموع مربعات خطا در مدل $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma X_{ijk} + \varepsilon_{ijk}$ در چه حالتی، حداکثر برابر مجموع مربعات خطای

مدل $y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}$ است؟

- (الف) فقط وقتی متغیر کمکی X از توزیع نرمال پیروی کند.
 (ب) فقط وقتی همبستگی متغیر کمکی X با متغیر پاسخ قوی باشد.
 (ج) فقط وقتی فرضیه $H_0: \gamma = 0$ رد شده باشد.
 (د) همواره

۸۱- اگر در رگرسیون لجستیک، برای یکی از متغیرهای مستقل دو حالتی داشته باشیم: $OR = 2.72$ و $SE(\hat{\beta}) = 0.4$

به دست آمده باشد فاصله اطمینان ۹۵ درصد برای نسبت شانس عبارت است از: (عدد نپر برابر ۲٫۷۲ است)

- (الف) (2.72 ± 0.8) (ب) (2.72 ± 0.4) (ج) $EXP(1 \pm 0.8)$ (د) $EXP(2.72 \pm 0.8)$

۸۲- فرض کنید متغیر X_2 را به مدل $Y = \beta_0 + \beta_1 X_1$ اضافه کنیم در این صورت شاخص $1 - \frac{SSE(X_1, X_2)}{SSE(X_2)}$ چه چیزی

را نشان می‌دهد؟

- (الف) ضریب تعیین جزئی بین X_1 و Y
 (ب) ضریب تعیین جزئی بین X_2 و Y
 (ج) کاهش نسبی تغییرات Y منتسب به X_1 در حالتی که X_2 در مدل حضور دارد
 (د) افزایش نسبی تغییرات Y منتسب به X_1 در حالتی که X_2 در مدل حضور دارد

۸۳- در یک طرح فاکتوریل 2^2 بدون تکرار پاسخ تیمارهای $A+B+$, $A-B+$, $A+B-$, $A-B-$ بترتیب اعداد ۵۰، ۲۰، ۲۰ و ۱۰ بدست آمده است. برآوردهای دو اثر متقابل $(AB)_{22}$ و $(AB)_{21}$ بترتیب از چپ به راست کدامند:
الف) $(-10, 10)$ ب) $(10, 10)$ ج) $(20, 20)$ د) $(-20, 20)$

۸۴- در یک طرح فاکتوریل 2^3 با سه عامل A , B , C اثر اصلی A با کدام یک از اثرات متقابل مخدوش (confound) می‌گردد:

الف) AC ب) AB ج) BC د) ABC

۸۵- در تحلیل طرح آزمایش‌های چند عاملی با استفاده از مدل‌های خطی، یک پیش فرض مهم مستقل بودن مشاهدات یا مستقل بودن جملات خطا می‌باشد. کدام یک از روش‌های آماری زیر می‌تواند این پیش فرض را به واقعیت نزدیک کرده و آن را معتبر سازد:

الف) Randomization
ب) Replication
ج) Blocking
د) Fractional Factorial Designing

۸۶- در طرح آزمایش یک عاملی با k گروه جهت آزمون همگنی واریانس‌ها از آزمون لون اصلاح شده (modified Leven's test) استفاده می‌شود. در این آزمون از کدامیک از تفاضل‌های زیر استفاده می‌گردد:
الف) تفاضل هر مشاهده با میانگین گروه خود.
ب) قدرمطلق تفاضل هر مشاهده با میانگین گروه خود.
ج) توان دوم تفاضل هر مشاهده با میانگین گروه خود.
د) قدرمطلق تفاضل هر مشاهده با میانه گروه خود.

۸۷- در یک طرح آزمایش نامتعادل با یک عامل سه سطحی، تعداد آزمایش‌ها در سطوح اول تا سوم بترتیب ۴ و ۹ و ۱۶ می‌باشند. اگر قرار باشد دو کنتراست متعامد از ترکیبات خطی سه میانگین μ_1 و μ_2 و μ_3 تشکیل شده و آزمون گردند، کدام جفت از ضرایب را برای کنتراست اول و دوم بترتیب از چپ به راست انتخاب می‌کنید:

الف) $(0.5, 0, -1)$ و $(0.5, 1, 1)$ ب) $(0, -1, 1)$ و $(-2, 1, 1)$

ج) $\left(0, -\frac{1}{9}, \frac{1}{16}\right)$ و $\left(-\frac{1}{4}, \frac{1}{9}, \frac{1}{16}\right)$ د) $(0, 1, 1)$ و $(2, -1, 1)$

۸۸- در آنالیز مولفه‌های اصلی، مقدار ویژه مرتبط با یک مؤلفه اصلی نشان‌دهنده کدام مورد است؟

الف) کوواریانس بین متغیرها
ب) واریانس توضیح داده شده توسط آن مؤلفه
ج) میانگین متغیرهای استاندارد شده
د) مجموع بارهای هر متغیر

۸۹- در تحلیل عاملی تأییدی (CFA)، کدام آماره معمولاً برای ارزیابی برازش کلی مدل استفاده می‌شود؟

الف) نمودار شن‌ریزه (Scree plot)

ب) آزمون کرویت بارتلت (Bartlett's sphericity test)

ج) جذر میانگین مربعات خطای تقریب (RMSEA)

د) معیار کیسر-میر-الکین KMO

۹۰- کدامیک از گزینه‌های زیر فرض معنی‌دار بودن ارتباط متغیر X_2 را با متغیر پاسخ در مدل زیر آزمون می‌کند؟

$$E(Y|X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_1 X_2$$

الف) $H_0 : \beta_2 = \beta_4$ ب) $H_0 : \beta_2 = 0$ ج) $H_0 : \beta_2 = \beta_4 = 0$ د) $H_0 : \beta_4 = 0$

۹۱- در خوشه‌بندی سلسله مراتبی، روش پیوند (Linkage) برای کدام یک از موارد زیر استفاده می‌شود؟

الف) تعیین فاصله بین خوشه‌ها

ب) تعیین تعداد خوشه‌ها

ج) تعیین واریانس بین خوشه‌ها

د) تعیین ارتباط بین خوشه‌ها

۹۲- روش تحلیل واریانس چند متغیره (MANOVA) برای داده‌های با اندازه‌گیری مکرر، در مقایسه با روش تحلیل واریانس تک متغیره (ANOVA):

الف) حتماً نیازمند فرض کرویت است.

ب) حتی در صورت نقض کرویت، قوی‌تر است.

ج) فقط با دو نقطه زمانی قابل استفاده است.

د) همبستگی درون آزمودنی را نادیده می‌گیرد.

۹۳- در کدامیک از روش‌های آماری چند متغیره زیر، انتخاب توزیع آماری برای متغیرها در همه مراحل و فرایندهای جانبی لازم نیست؟

الف) خوشه‌ای (Cluster analysis)

ب) مولفه‌های اصلی (Principal component analysis)

ج) همبستگی کانونی (Canonical correlation analysis)

د) تحلیل ممیزی (Discriminant analysis)

۹۴- کدامیک از روش‌های زیر مکمل همدیگر هستند؟

الف) تحلیل عاملی و تحلیل مولفه‌های اصلی

ب) تحلیل خوشه‌ای و تحلیل عاملی

ج) تحلیل همبستگی کانونی و رگرسیون چندمتغیره

د) تحلیل ممیزی و تحلیل خوشه‌ای

۹۵- در یک مدل رگرسیون لجستیک رتبه‌ای با بخت‌های متناسب (proportional odds)، ضریب مثبت برای یک متغیر پیش‌بینی‌کننده چه تفسیری دارد؟

الف) احتمال قرار گرفتن در پایین‌ترین سطح پاسخ افزایش می‌یابد.

ب) احتمال قرار گرفتن در سطوح بالاتر پاسخ کاهش می‌یابد.

ج) نسبت شانس تجمعی برای قرار گرفتن در سطوح پایین‌تر پاسخ نسبت به سطوح بالاتر افزایش می‌یابد.

د) نسبت شانس تجمعی برای قرار گرفتن در سطوح بالاتر پاسخ نسبت به سطوح پایین‌تر افزایش می‌یابد.

۹۶- در یک جدول توافقی سه‌بعدی شامل متغیرهای A، B و C، پژوهشگر مدلی با اثرات متقابل دوتایی AB و AC برازش داده است. این مدل کدام یک از روابط زیر را بررسی می‌کند؟

الف) A و B مستقل از C هستند.

ب) B و C به شرط A مستقل از هم هستند.

ج) A و B به شرط C مستقل از هم هستند.

د) اثر متقابل سه‌گانه بین A و B و C وجود دارد.

۹۷- در رگرسیون خطی چندگانه $E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$ اگر ضرایب همبستگی Y با X_1 و X_2 به ترتیب ۰,۳ و ۰,۶ باشد چند درصد تغییرات Y توسط مدل بالا بیان می‌شود؟

الف) ۰,۳۶ (ب) ۰,۴۵ (ج) بین ۰,۳۶ و ۰,۴۵ (د) بین ۰,۰۹ و ۰,۴۵

۹۸- در برازش مدل $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$ در یک نمونه ۲۳ نفری، مقادیر $SSE=20$ و $SSR=100$ به دست آمده است. کدام یک از فرضیه‌های زیر در سطح خطای نوع اول ۰,۰۵ حتما رد می‌شود؟

الف) $H_0: \beta_0 = \beta_1 = \beta_2 = 0$

ب) $H_0: \beta_1 = \beta_2 = 0$

ج) هر دو $H_0: \beta_0 = 0$ و $H_0: \beta_2 = 0$

د) هر دو $H_0: \beta_0 = 0$ و $H_0: \beta_2 = 0$

۹۹- در رگرسیون لجستیک $\log\left(\frac{p}{1-p}\right) = -100 + 2.5X$ که در آن X غلظت یک حشره کش (میلی گرم) و p احتمال مرگ حشرات باشد، کدام غلظت می‌تواند ۵۰ درصد کشندگی ایجاد کند.

الف) ۴۰ میلی گرم (ب) ۲,۵ میلی گرم (ج) ۱۰۰ میلی گرم (د) ۵۰ میلی گرم

۱۰۰- اگر نسبت دو متغیر برای افراد مختلف جامعه ثابت باشد آنگاه:

الف) بین دو متغیر قطعاً ارتباط خطی وجود ندارد.

ب) بین دو متغیر ممکن است ارتباط خطی وجود نداشته باشد.

ج) بین دو متغیر ارتباط خطی کامل وجود دارد ولی خط از مبدا عبور نمی‌کند.

د) بین دو متغیر ارتباط خطی کامل وجود دارد و خط از مبدا عبور می‌کند.

زبان تخصصی و عمومی

زبان تخصصی

101- If the occurrence of event A is incompatible with the occurrence of event B, these two events:

- a) are independent of each other
- b) are associated with each other
- c) have intersection
- d) are mutually exclusive

102- In a simple linear regression analysis that examines the relationship between physical activity time and blood pressure, physical activity time is:

- a) an adjusting variable
- b) an exploratory variable
- c) an explanatory variable
- d) a confounding variable

103- If a sample is divided into q distinct sections with equal frequencies, each section is called a:

- a) quartile
- b) quantile
- c) quintile
- d) quadrant

104- Data that are described as 'heterogeneous' are:

- a) exactly the same as each other
- b) relatively similar to each other
- c) relatively dissimilar to each other
- d) significantly associated

105- Provincial accidental fatality rates are going to be compared before and after the enactment of seat-belt laws in a country. The study design is a:

- a) Randomized controlled trial
- b) Quasi experimental
- c) Randomized block design
- d) Factorial design

106- The number needed to treat (NNT) of a new treatment compared with that of an existing treatment is:

- a) the number of patients need to receive the new treatment instead of the existing treatment.
- b) the number of patients need to receive the new treatment, in order to produce one additional desirable outcome.
- c) the number of patients who have received placebo instead of the new treatment.
- d) the number of patients who have received the new treatment but had adverse outcome.

107- Which studies provide a firm base for the establishment of cause and effect relationship between one or more factors and a response variable?

- a) Observational studies
- b) Experimental studies
- c) Ecological studies
- d) Retrospective studies

108- Which is correct for a randomization process?

- a) Selecting simple random samples from populations in a cross sectional study
- b) Selecting random samples from each stratum of the population
- c) Selecting random samples from sub-populations, separately, for each experimental group
- d) Assigning treatments randomly to the experimental units

109- In the theory of testing hypotheses, an acceptance region is:

- a) a region in the sample space.
- b) a region out of the sample space.
- c) independent of the sample space.
- d) wider than the sample space.

110- Bartlett's test is an approximate test for the homogeneity of a set of

- a) means of several independent Normal samples.
- b) means of several dependent Normal samples.
- c) variances of several independent Normal samples.
- d) variances of several dependent Normal samples.

111- Bayesian method is a form of inference which regards parameters as being

- a) fixed values.
- b) fixed values with prior distribution.
- c) random variables possessed of prior distributions.
- d) random parameters with fixed values.

112- Bernoulli trials are sequences of events in which successive trials are

- a) independent with constant probability of success.
- b) independent with random probability of success.
- c) dependent with constant probability of success.
- d) dependent with random probability of success.

113- Hyper-geometric distribution is a distribution of discrete variable that is associated with sampling from

- a) a finite population without replacement.
- b) an infinite population without replacement.
- c) an infinite population with replacement.
- d) a finite population with replacement.

114- Kurtosis describes the extent of the relative steepness of ascent in the neighborhood of the

- a) mean.
- b) median.
- c) mode.
- d) first quartile.

115- Bias and random error distort the statistical results, however

- a) both are balanced out on the average.
- b) bias is balanced out on the average but random error is not.
- c) random error is balanced out on the average but bias is not.
- d) both are not balanced out on the average.

■ **Part one: Reading comprehension**

Read the following passages carefully and then answer the questions that follow. Base your answers on the information in the passage only.

Passage 1:

Artificial Intelligence (AI) is revolutionizing healthcare by enhancing diagnostic accuracy, personalizing treatment plans, and streamlining administrative tasks. In diagnostics, AI algorithms analyze medical images such as X-rays and MRIs with remarkable precision, often surpassing human radiologists in detecting anomalies like tumors or fractures. For instance, machine learning models trained on vast datasets can identify early signs of diseases such as cancer or Alzheimer's, enabling timely interventions that improve patient outcomes. Moreover, AI-powered predictive analytics forecast disease outbreaks and patient deteriorations, allowing healthcare providers to allocate resources more efficiently. In treatment personalization, AI systems process genetic information and medical histories to recommend tailored therapies, reducing the trial-and-error approach in prescribing medications. This not only minimizes adverse effects but also optimizes recovery times. Administrative applications of AI include automating scheduling, billing, and electronic health record management, which reduces paperwork and frees up medical staff to focus on patient care. However, the integration of AI in healthcare raises ethical concerns, including data privacy, algorithmic bias, and the potential displacement of jobs. Ensuring that AI systems are trained on diverse datasets is crucial to avoid biases that could disproportionately affect minority groups. Additionally, robust cyber-security measures are essential to protect sensitive patient information from breaches. Despite these challenges, the benefits of AI in healthcare are substantial, promising a future where medical care is more accessible, efficient, and effective. Governments and organizations are investing heavily in AI research to harness its potential while addressing the associated risks. As AI continues to evolve, its role in healthcare will likely expand, incorporating advanced technologies like natural language processing for better doctor-patient communication and robotics for surgical precision. Ultimately, AI has the potential to democratize healthcare, making high-quality medical services available to underserved populations around the world.

116- It is stated that artificial intelligence in diagnostics.

- a) outperforms human experts in finding masses and broken bones
- b) helps reduce the workload of administrative staff
- c) assists in tailoring treatment plans to individual patients
- d) forecasts abnormal bone fractures and cancerous tumors

117- As regards the predictive models, AI can help practitioners

- a) enhance patient deterioration and disease epidemics
- b) train hospital databases and warn about diseases
- c) take well-timed actions to treat patients
- d) enable datasets to identify initial signs of diseases

118- In personalized treatment, artificial intelligence

- a) helps doctors pile up patient records
- b) minimizes the accuracy of hospital billing systems
- c) reduces guesswork in ordering proper medication
- d) prevents routine access to patient records

119- The ethical concern to avoid affecting minority groups requires

- a) taking security measures to safeguard patient communication
- b) using varied datasets to help prevent bias algorithms
- c) protecting such groups against job displacement
- d) undermining healthcare issues proportionately

120- AI helps the democratization of healthcare by

- a) improving doctor-patient communication through speech recognition tools
- b) eliminating the need for government supervision on AI development
- c) reducing the cost of robotic equipment used in potential surgeries
- d) making advanced medical services more accessible to deprived communities

Passage 2:

Health inequality refers to the systematic differences in health outcomes and access to healthcare services among different population groups, often influenced by socioeconomic status, race, geography, and education. In many countries, individuals from lower-income backgrounds experience higher rates of chronic diseases such as diabetes, heart disease, and obesity due to limited access to nutritious food, safe environments for physical activity, and preventive medical care. For example, urban-rural divides exacerbate these disparities, with rural populations facing shortages of healthcare providers and longer travel times to facilities, leading to delayed diagnoses and poorer prognoses. Racial and ethnic minorities frequently encounter barriers like discrimination within healthcare systems, cultural insensitivity, and language obstacles, which contribute to lower vaccination rates and higher maternal mortality. Education plays a pivotal role; those with higher education levels are more likely to understand health information, adhere to treatment regimens, and seek timely care, whereas less educated individuals may misunderstand symptoms or avoid seeking help due to stigma. Economic factors compound these issues, as low-wage workers often lack health insurance or cannot afford out-of-pocket expenses, resulting in untreated conditions that worsen over time. Governments and organizations are attempting to address health inequality through policies like expanding public health insurance, investing in community health centers, and promoting health education campaigns targeted at vulnerable groups. However, challenges persist, including political resistance to funding increases and the slow pace of systemic change. Technological advancements, such as telemedicine, offer promise in bridging geographical gaps, but digital divides—where low-income or elderly populations lack internet access—can inadvertently widen inequalities. Ultimately, reducing health inequality requires a multifaceted approach involving policy reform, community engagement, and equitable resource distribution to ensure that all individuals, regardless of background, can achieve optimal health outcomes and live longer, healthier lives.

121- This reading passage mainly discusses

- a) the roles of technology in healthcare
- b) impacts of urban living on healthcare
- c) advantages of higher education levels in health
- d) causes of and solutions to health disparities

122- It is stated in the passage that health inequality is influenced by

- a) disease stigma
- b) socioeconomic condition
- c) elderly access to telemedicine
- d) political resistance

123- Governments reduce health inequality by

- a) educating individuals with equitable resources
- b) increasing out of pocket expenses
- c) adhering to treatment regimens
- d) expanding community health facilities

124- Which of the following is NOT mentioned as an obstacle to racial minorities?

- a) Cultural insensitivity
- b) Communication issues
- c) Understanding health information
- d) Inequality of health care services

125- It is inferred that technological advancements are hindered by

- a) digital divides
- b) equitable resource
- c) internet access
- d) policy reform





■ **Part two: Vocabulary**

Complete the following sentences, choosing the most appropriate answer (a, b, c, or d).

126- The medication the virus from reproducing in the body.

- a) generates
- b) inhibits
- c) scatters
- d) radiates

127- His appendix leading to emergency surgery.

- a) ruptured
- b) nurtured
- c) alleviated
- d) ameliorated

128- Doctors often treatment plans to each patient's unique health profile.

- a) predispose
- b) generalize
- c) customize
- d) expose

129- Evidence-based research modern medical practices which, in turn, strengthens the efficacy of treatment plans.

- a) depreciates
- b) undermines
- c) debilitates
- d) underpins

130- Elderly patients are falls due to physical weakness, so they need extra care and attention.

- a) content with
- b) amazed with
- c) adjacent to
- d) prone to

موفق باشید





بسمه تعالی

دبیرخانه شورای آموزش علوم پایه پزشکی، بهداشت و تخصصی و مرکز سنجش آموزش پزشکی با هدف ارتقای کیفیت سوالات و بهبود روند اجرای آزمون‌ها، پذیرای درخواست‌های بررسی سوالاتی است که در قالب مشخص شده زیر از طریق اینترنت ارسال می‌گردد، تا کار رسیدگی با سرعت و دقت بیشتری انجام گیرد.

ضمن تشکر از همکاری داوطلبان محترم موارد ذیل را به اطلاع می‌رساند:

- ۱- کلید اولیه سوالات ساعت ۱۴ مورخ ۱۴۰۴/۰۸/۱۱ از طریق سایت اینترنتی www.sanjeshp.ir اعلام خواهد شد.
- ۲- اعتراضات خود را از ساعت ۱۴ مورخ ۱۴۰۴/۰۸/۱۱ لغایت ساعت ۸ صبح مورخ ۱۴۰۴/۰۸/۱۷ به آدرس اینترنتی بالا ارسال نمایید.
- ۳- اعتراضاتی که به هر شکل خارج از فرم ارائه شده، بعد از زمان تعیین شده و یا به صورت غیراینترنتی (حضوری) ارسال شود، مورد رسیدگی قرار نخواهد گرفت.

تذکر مهم:

- * فقط اعتراضات ارسالی در فرصت زمانی تعیین شده، مورد بررسی قرار گرفته و پس از تاریخ مذکور به هیچ عنوان ترتیب اثر داده نخواهد شد.
- * از تکرار اعتراضات خود به یک سوال پرهیز نمایید. تعداد اعتراض ارسالی برای یک سوال، ملاک بررسی نمی‌باشد و به کلیه اعتراضات ارسالی اعم از یک مورد و یا بیشتر رسیدگی خواهد شد.

دبیرخانه شورای آموزش علوم پایه
پزشکی، بهداشت و تخصصی

مرکز سنجش آموزش علوم پزشکی

نام:		نام خانوادگی:		کد ملی:	
نام رشته:		نام درس:		شماره سوال:	
نام منبع معتبر		سال انتشار		صفحه	
				پاراگراف	
				سطر	

سوال مورد بررسی:

- ☐ بیش از یک جواب صحیح دارد. (با ذکر جواب‌های صحیح)
- ☐ جواب صحیح ندارد.
- ☐ متن سوال صحیح نیست.

توضیحات: